

葛兴统

✉ xingtong.ge@gmail.com · 📞 (+86) 18810927898 · 🌐 XingtongGe · 📅 2001.09.27

教育背景

北京理工大学 硕士，软件工程 2022年9月 – 2024年6月

北京理工大学 学士，计算机科学与技术 2018年9月 – 2022年6月

- GPA: 87.5/100
- 毕业论文：面向机器视觉任务的图像压缩编码算法研究（北京理工大学优秀毕业论文，前5%）

实习 & 科研经历

商汤科技研究院 大模型服务部门 见习算法研究员 2024年1月 – 至今

- 参与商汤“秒画”团队研发，使用生成模型对人脸自动上妆进行调优，解决部分图片上妆效果崩坏等问题，相关功能上线“秒画合拍”小程序
- 参与文生图模型研发，引入语义先验改进扩散生成模型隐空间分布，改善模型生成图像风格及结构多样性
- 改进文生图大模型 Latent Consistency Model 训练策略，实现图像生成质量优化

商汤科技研究院 ISP&Codec 部门 AI 视频编解码实习生 2023年5月 – 2024年1月

- 主要研究课题为基于深度学习的图像与视频压缩编码
- 基于深度学习的视觉可控通用视频压缩 (CVPR 2024)
 - 提出一种视觉可控的深度学习视频压缩方法，在编码阶段控制编码器适应于下游机器视觉任务（多目标追踪、视频目标检测等）或人眼视觉重建，简化多任务下的视频编解码流程；
 - 提出采用编码模式预测策略，对原始编码预测帧进行面向机器视觉优化，有效减少编码所需码率，同时进一步引入编码 GoP 结构预测策略，实现解码视频对下游任务超过 25% 码率-准确率性能提升。
- 基于隐式神经网络的视频表征与压缩 (CVPR 2024)
 - 在隐式神经网络的视频表征中加入时序条件先验等多种增强方法，提出一种通用的视频表征和压缩增强框架，增强现有多种视频表征模型性能

上海交通大学 & 北京理工大学 科研学生 2021年10月 – 至今

- 面向机器视觉的前处理图像压缩 (IEEE TCSVT submitted)
 - 在传统图像编解码器基础上提出一种前处理增强的图像编解码框架，在相同下游机器视觉任务性能下相对增强前节省超过 20% 码率，在不同任务之间具有强泛化性；
 - 解决传统编码器不可微分问题，引入可微分代理网络近似拟合传统编解码器，代理实现端到端联合优化

项目经历

3D/2D Gaussian Splatting 2023年12月 – 2024年3月

- CLIP Guided 3D Gaussian Splatting: 联合 OpenCLIP 语义表征与 3D GS，实现符合语义的 3D 物体生成
- 2D Gaussian Splatting 图像表征与应用 (ECCV 2024 submitted): 提出 2D GS，首个将 GS 技术应用于图像表征与压缩，压缩表现超越现有基于 INR 方法，在流行数据集上实现超过 1000FPS 解码速度

基于深度学习的低光图像增强算法研究与改进 团队第一负责人 2022年10月-2022年12月

- 对来自 CVPR 的一种无参照低光图像增强算法进行代码实现，并在该模型基础上提出一种改善色彩增强能力的低光图像增强方法，成功实现在多种图像增强测试集下 PSNR、SSIM 等多个指标的图像增强性能改进

2021 NAIC “华为昇腾杯” AI+ 视觉特征编码大赛 团队第一负责人 2021年12月 – 2022年3月

- 提出一种将图像压缩任务与行人车辆重识别任务相结合的联合优化方法，在大赛数据集上进行图像的特征提取、特征压缩重建、行人车辆重识别，最终得分排名前 30 名（总参赛队伍 1924 支）

学术论文

- Ge X, Luo J, Zhang X, Xu T, et al. Task-Aware Encoder Control for Deep Video Compression. (CVPR 2024)
- Zhang X*, Ge X*, Xu T, He D, et al. GaussianImage: 1000 FPS Image Representation and Compression by 2D Gaussian Splatting. (ECCV 2024 under review)
- Zhang X, Yang R, He D, Ge X, et al. Boosting Neural Representations for Videos with a Conditional Decoder. (CVPR 2024)
- Lu G, Ge X, Zhong T, et al. Preprocessing Enhanced Image Compression for Machine Vision[J]. arXiv preprint arXiv:2206.05650, 2022. (IEEE TCSVT under review)